Non-invasive embryo assessment in IVF

Vergouw, C.G.

2014

document version
Publisher's PDF, also known as Version of record


Chapter
6

NO EVIDENCE THAT EMBRYO SELECTION BY NEAR-INFRARED SPECTROSCOPY IN ADDITION TO MORPHOLOGY IS ABLE TO IMPROVE LIVE BIRTH RATES: RESULTS FROM AN INDIVIDUAL PATIENT DATA META-ANALYSIS

C.G. Vergouw
M.W. Heymans
T. Hardarson
I.A. Sfontouris
K.A. Economou
A. Ahlström
L. Rogberg
T.G. Lainas
D. Sakkas
D.C. Kieslinger
E.H. Kostelijk
P.G.A. Hompes
R. Schats
C.B. Lambalk

Human Reproduction: accepted

ABSTRACT

Study question: What is the value of embryo selection by metabolomic profiling of culture medium with near-infrared (NIR) spectroscopy as an adjunct to morphology, compared with embryo selection by morphology alone, based on an individual patient data meta-analysis (IPD-MA)?

Summary answer: The IPD-MA indicates that the live birth rate after embryo selection by NIR spectroscopy and morphology is not significantly different compared with the live birth rate after embryo selection by morphology alone.

What is known already: Retrospective proof of principle studies have consistently shown that high NIR viability scores are correlated with a high implantation potential of embryos. However, randomized controlled trials (RCTs) have generally shown no benefit of the NIR technology over embryo morphology, although there have been some conflicting results between pregnancy outcomes on different days of embryo transfer.

Study design, size, duration: This IPD-MA included all existing RCTs (n=4) in which embryo selection by morphology was compared with embryo selection by morphology and the use of NIR spectroscopy of spent embryo culture medium by the ViaMetrics-E™.

Participants/materials, setting, methods: Searches of PubMed, the Cochrane Library and the WHO International Clinical Trials Registry were conducted and the sole manufacturer of the ViaMetrics-E™ was consulted to identify clinics where an RCT comparing embryo selection by morphology with embryo selection by morphology and the use of the ViaMetrics-E™ (NIR viability score) was performed. A total of 20 citations were potentially eligible for inclusion, two of which met the inclusion criteria. The manufacturer of the ViaMetrics-E™ provided two additional clinical sites of use. In total, four RCTs were identified as eligible for inclusion. The IPD-MA was based on a fixed effect model due to the lack of heterogeneity between included studies. Differences between study groups were tested and reported using logistic regression models adjusted for significant confounders. The pooled analysis of the primary outcome led to a total sample size of 924 patients: 484 patients in the control group (embryo selection by morphology alone) and 440 patients in the treatment group (embryo selection by morphology plus NIR spectroscopy).

Main results and role of chance: The live birth rates in the control group and the NIR group were 34.7% (168 of 484) and 33.2% (146 of 440), respectively. The pooled odds ratio (OR) was 0.98 (95% confidence interval [CI] 0.74-1.29), indicating no difference in live birth rates between the two study groups. The data of the four studies showed no

significant heterogeneity (I^2=26.2% P=0.26). The multivariate regression analysis including all confounders showed that the study group (i.e. embryo selection by morphology or embryo selection by morphology plus NIR) was not related to live birth (OR 0.97, 95% CI 0.73-1.29).

Limitations and reasons for caution: The availability of at least two similar best quality embryos as an inclusion criterion prior to transfer in the two largest RCTs might have caused a selection bias towards a better prognosis patient group.

Wider implications of the findings: There is at present no evidence that NIR spectroscopy of spent embryo culture media in its current form can be used in daily practice to improve live birth rates.

INTRODUCTION

Although in many countries the implementation of single embryo transfer (SET) has been strictly put into practice, the concept of simultaneous transfer of multiple embryos in order to maximize pregnancy results per IVF cycle still persists. This strategy has led to an increased incidence of multiple pregnancies and their associated pregnancy-related health problems and complications for neonates[1-3]. To aid in the selection of one embryo for transfer, several strategies have been examined in the search for additional markers of embryo viability to supplement current criteria for embryo selection. One of these new technologies is metabolomic profiling of spent embryo culture media with the use of near-infrared (NIR) spectroscopy.

Several proof of principle studies have suggested the use of the NIR and the derivation of a viability score as a tool to assess embryo viability[4-12]. The viability score is generated by comparing regions within NIR spectral profiles of spent embryo culture medium that discriminate between implanted and non-implanted embryos. Quantification and subsequent expression as a viability score of spectral profiles is done by multivariate algorithms. All of the above mentioned proof of principle studies have consistently reported that high NIR viability scores are correlated with high implantation potential of embryos, and that this correlation is independent of embryo morphology.

After the promising results from retrospective studies, randomized controlled trials (RCTs) were initiated to prospectively determine whether embryo selection by metabolomic profiling of culture medium with NIR spectroscopy in adjunct to morphology could improve ongoing pregnancy and live birth rates compared with embryo selection by morphology alone. Hardarson et al.[13] were the first to report the outcome of an interim analysis from a prospective RCT evaluating day 2 or day 5 embryo culture medium by NIR with ongoing pregnancy as primary outcome. This interim analysis did not reveal a significant difference for the combined day 2 and day 5 data with regard to ongoing pregnancy rates, when an embryo was selected by both morphology and the NIR viability score compared with embryo selection by morphology alone. However, in the day 2 subgroup, a minor, but not significant improvement of 4.5% in ongoing pregnancy rates was seen in favour of the morphology and NIR viability score group compared with the morphology only group. On the other hand, in the day 5 subgroup, morphology alone was 6.5% better. Vergouw et al.[14] showed in a double blind RCT that day 3 embryo selection by metabolomic profiling of culture medium with NIR spectroscopy as an addition to morphology was not able to improve ongoing pregnancy and live birth rates compared with embryo selection by morphology alone. Sfontouris et al.[15], reported significant higher ongoing implantation

rates for day 5 transfers (but not for day 2 and day 3 transfers) yet similar ongoing pregnancy and live birth rates from their RCT in which embryo selection by metabolomic profiling with NIR spectroscopy plus morphology was compared with morphology alone. However, this study lacked adequate power, because the study was terminated early due to the voluntary market withdrawal of the NIR instrument for commercial reasons.

The differences in outcomes between the reported randomized controlled trials led us to perform an individual patient data meta-analysis (IPD-MA). As an alternative to a conventional meta-analysis, an IPD-MA is thought to provide a more reliable estimate of treatment effect[16]. The aim of this study was to use IPD-MA to assess the effectiveness of metabolomic profiling of culture medium with NIR spectroscopy technology (NIR viability score) as an adjunct to embryo morphology in IVF.

MATERIALS AND METHODS

SELECTION OF STUDIES AND DATA EXTRACTION

Studies that reported the results of an RCT comparing embryo selection by morphology with embryo selection by morphology and the use of NIR spectroscopy of spent embryo culture medium by the ViaMetrics-E™ were identified for inclusion. First of all, this was done by contacting the (sole) manufacturer of the ViaMetrics-E™, who provided (after consent) the names of researchers that had performed such an RCT. For completeness, a literature search was performed in PubMed, the Cochrane Library and the WHO International Clinical Trials Registry. We used the following terms: 'randomized controlled trials' or 'systematic reviews' or 'meta-analysis', 'metabolomic profiling' or 'viability score' and 'in vitro fertilization'.

We selected studies with a randomized design with a population of women who underwent IVF or ICSI treatment and where the embryo selection was performed by morphology compared with embryo selection by morphology plus NIR technology. The study protocol was approved by the institutional review board.

DATA ACQUISITION

First authors of the eligible studies were contacted by email. We asked whether they were willing to share their data for this project. Authors who agreed to participate were asked to send the complete, original, anonymous data set. The data set was accepted in all formats, providing that variables and categories were adequately labelled. Data about the following parameters were requested: anonymous patient

identifiers, allocation of patient, maternal age, maternal BMI, parity, duration of infertility, medical cause of infertility, number of previous IVF attempts, treatment type (type of pituitary regulation; dosage and type of ovarian stimulation), number of oocytes at ovum pick-up (OPU), fertilization method (IVF or ICSI), number of fertilized oocytes, number of (good quality) embryos and number of embryos transferred, day of embryo transfer, if applicable: the viability score of the transferred embryo, number of embryos cryopreserved, number of fetal cardiac activity (FCA) at 12 weeks of gestational age and finally, the number of babies born alive. Authors were also asked to provide the procedure of the randomization and concealment of allocation.

ASSESSMENT OF STUDY QUALITY

Study quality of the selected studies was independently assessed by two authors (CGV and CBL). For this, the Cochrane checklist for evaluation of RCTs[17] was used. The validity of a trial that did not correspond to the standards of the Cochrane checklist was considered compromised and could be excluded depending on the intensity to which the study was compromised.

Completeness of the databases was checked based on the available data of the minimal data requirements that was sent by the authors. When data were missing or discordant, authors were approached for clarification. There was no strict cut-off set for exclusion of the study when data was missing.

STATISTICS

The primary outcome of this study was the live birth rate per couple. Descriptive statistics were used to explore the characteristics of the studies and the between study differences. The data of all studies were pooled and an intention-to-treat (ITT) analysis was conducted. Statistical heterogeneity between trials was explored and tested by the I^2 statistic in a random effects model. Variation between trials was also tested using a two-level logistic mixed model, where patients will correspond to the first level and study will correspond to the second level. In these models, we explored whether the heterogeneity in treatment and patient characteristics needed to be taken into account by including random centre effects. We tested in the multilevel mixed model if the treatment effect differed between hospitals by testing if the slope effect (of treatment) differed significantly from zero. This was not the case. This is an indication that the heterogeneity between hospitals in treatment effect was low. Therefore, a logistic regression model was justified for our data.

Because all heterogeneity tests showed that between-studies variability was low and did not influence effects estimates, differences between study groups in the current

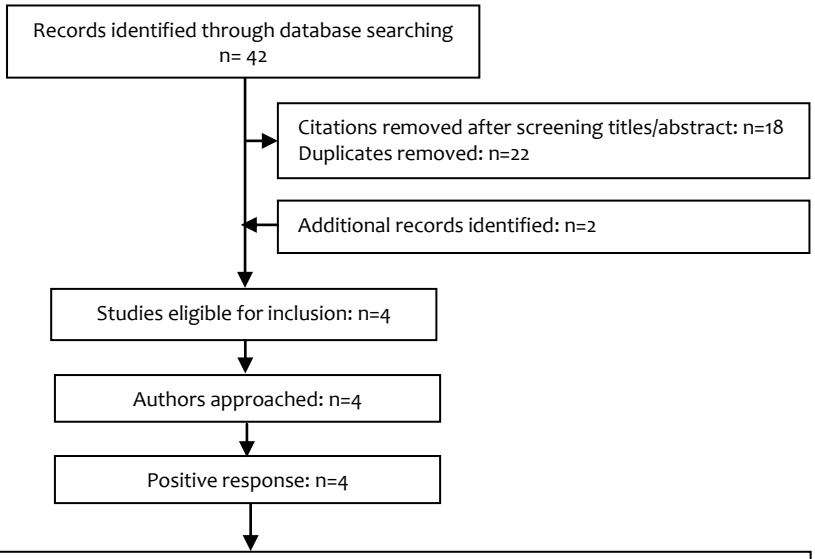
study were finally tested and reported using a one-level logistic regression model. In this one-level regression model, confounding was tested by determining the change in the treatment effect of the group variable, when a potential confounding variable was added to the model. A strong confounding effect was defined as a change of >10% on the treatment group variable when the confounding variable was added to the model. In this way, the strongest confounding variables were identified and in the case of a strong influence on the treatment estimate, these were adjusted for in the regression models. Potential confounding variables assessed at baseline were: maternal age, maternal BMI, parity, duration of infertility, medical cause of infertility, number of previous IVF attempts, treatment type, number of oocytes at OPU, fertilization method, number of fertilized oocytes, number of (good quality) embryos, number of embryos transferred and the day of embryo transfer. We also tested for interaction effects of the treatment group with all potential confounders. Furthermore, all models were adjusted for trial ID. Data were analysed using SAS Proc Glimmix (SAS Institute Inc., Cary, NC, USA), SPSS 20.0 (SPSS Inc., Chicago, IL, USA) and STATA 12 (Statacorp LP, College Station, TX, USA).

RESULTS

An overview of the selection of the trials and their methodological quality is presented in Figure 1. The literature search produced a total of 42 citations and after the removal of duplicates, 20 citations were potentially eligible. After detailed evaluation, we identified two primary articles that met the inclusion criteria. The manufacturer of the ViaMetrics-E™ provided two additional contacts of sites that conducted an RCT comparing embryo selection by morphology with embryo selection by morphology and the use of the ViaMetrics-E™ (NIR viability score). In total, four RCTs were identified as eligible for inclusion in this IPD-MA. All authors gave their consent to participate in this study and we obtained four databases of which three had been used for previous publications[13-15]. The two larger RCTs were performed using prototype instruments while the smaller RCTs were performed using initial commercial instruments.

The study size of the original clinical trials varied between 55 and 417 (ITT population) and the four databases combined contained data of 924 IVF/ICSI cycles of 924 patients. A total of 484 patients had been randomized in the control group, where embryos for transfer were selected by morphology alone, and 440 patients had been randomized in the treatment group, where embryos were selected by morphology plus NIR spectroscopy. Baseline characteristics of the 924 women and subsequent IVF/ICSI cycles are presented in Table 1. Data on female age, BMI, medical cause of infertility,

number of previous IVF attempts, total dosage of gonadotropins administered, fertilization method, number of oocytes retrieved, number of fertilized oocytes, number of good quality embryos available for transfer, number of embryos transferred and day of transfer were available from all studies.



Included studies: n=4

Trial	Method of randomization and allocation of concealment	Blinding	Type of analysis in original article	End-point
Hardarson et al.[13]	Computerized tables Adequate concealment by sealed envelopes	Double blind	Intention-to-treat	Live birth
Vergouw et al.[14]	Computerized tables Adequate concealment by sealed envelopes	Double blind	Intention-to-treat	Live birth
Sfontouris et al.[15]	Computerized tables No concealment	Not blinded	Intention-to-treat	Live birth
Economou et al.	Computerized tables Adequate concealment by sealed envelopes	Double blind	N/A	Live birth

Figure 1. Flowchart of included studies.

Table 1. Baseline characteristics of studies that provided IPD.

	Hardarson	Vergouw	Sfontouris	Economou	Total
Total no. of patients	327	417	125	55	924
Maternal age (years)	35.5 ± 4.2	34.2 ± 4.3	35.2 ± 4.4	33.4 ± 3.1	34.8 ± 4.2
Maternal BMI	23.6 ± 3.8 n=315	24.5 ± 4.5 n=413	24.7 ± 4.6	23.2 ± 3.7	24.1 ± 4.3 n=908
Primary infertility	N/A	241 (57.8)	N/A	48 (87.3)	289 (61.2) n=472
Duration of infertility (years)	N/A	3.2 ± 2.0	3.9 ± 3.2	2.6 ± 1.7	3.3 ± 2.3
Medical cause of infertility					
Male	103 (31.5)	202 (48.4)	55 (44.0)	11 (20.0)	371 (40.2)
Tubal	33 (10.1)	62 (14.9)	34 (27.2)	10 (18.2)	139 (15.0)
Unexplained	114 (34.9)	75 (18.0)	11 (8.8)	18 (32.7)	218 (23.6)
Endometriosis	N/A	35 (8.4)	7 (5.6)	1 (1.8)	43 (4.7)
Other	77 (23.5)	43 (10.3)	18 (14.4)	15 (27.3)	153 (16.6)
No. of previous IVF attempts	0.6 ± 1.0	0.3 ± 0.6	1.4 ± 1.9	0.6 ± 1.2	0.6 ± 1.0
Total dosage of gonodatropins administered (IU)	2009 ± 794	2158 ± 899	2467 ± 1045 n=120	1438 ± 745	2103 ± 901 n=919
Fertilization method					
IVF	188 (57.5)	216 (51.8)	20 (16.0)	0	424 (45.9)
ICSI	139 (42.5)	201 (48.2)	78 (62.4)	45 (81.8)	463 (50.1)
IVF and ICSI	0	0	27 (21.6)	10 (18.2)	37 (4.0)
No. of oocytes retrieved	9.1 ± 5.0	11.3 ± 6.7	14.9 ± 6.7	13.5 ± 4.5	11.1 ± 6.4
No. of fertilized oocytes	5.9 ± 3.4	6.5 ± 4.4	8.8 ± 4.1	8.8 ± 2.9	6.7 ± 4.1
No. of good quality embryos available for transfer	2.9 ± 1.2	3.9 ± 3.2	2.2 ± 1.7	4.6 ± 2.6	3.4 ± 2.5
No. of embryos transferred	1	1	2.9 ± 0.7	2.8 ± 0.4	1.4 ± 0.8
Day of transfer	day 2 (n=170) day 5 (n=157)	day 3 (n=417)	day 2 (n=35) day 3 (n=40) day 5 (n=50)	day 3 (n=55)	day 2 (n=205) day 3 (n=512) day 5 (n=207)
No. of embryos cryopreserved	N/A	3.7 ± 3.4	1.0 ± 2.2	2.2 ± 1.7	3.0 ± 3.3 n=597

Data are mean ± SD or n (%).

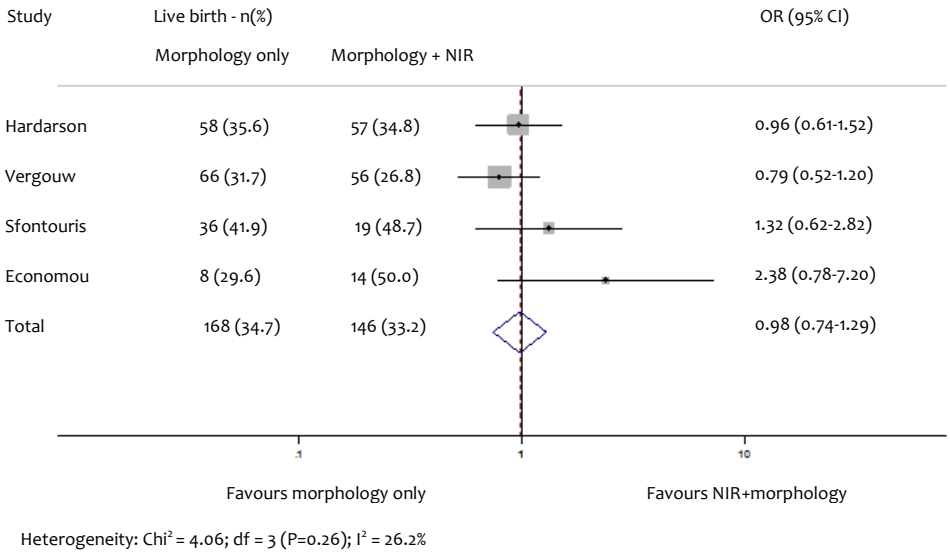


Figure 2. Forest plot for the results of the meta-analysis with IPD.

Three studies provided data on the duration of infertility and number of embryos cryopreserved and two studies provided data on parity.

Data on the primary end-point live birth was available for all 924 patients. The live birth rate in the control group was 34.7% (168 of 484) and the live birth rate in the NIR group was 33.2% (146 of 440). Figure 2 shows the results of the IPD-MA logistic regression analysis. The pooled odds ratio (OR) was 0.98 (95% confidence interval [CI] 0.74-1.29), indicating no differences in live birth rates between the two study groups. The data of the four studies showed no significant heterogeneity (I^2=26.2% P=0.26). Therefore, we applied a fixed effect model in this IPD-MA.

For model building and the evaluation of potential confounders, we used the variables that were provided by all four study sites (n=924; i.e. maternal age, medical cause of infertility, number of previous IVF attempts, treatment type, number of oocytes at OPU, fertilization method, number of fertilized oocytes, number of good quality embryos, number of embryos transferred and the day of embryo transfer). The test for interaction effects of the treatment group with all potential confounding variables was not significant. Logistic regression analyses identified maternal age, medical cause of infertility, number of previous IVF attempts, number of oocytes at OPU and the number of fertilized oocytes as the strongest confounding factors. The multivariable logistic regression analysis including these confounding variables showed that study group (i.e.

embryo selection by morphology or embryo selection by morphology plus NIR) was not related to live birth (OR 0.97, 95% CI 0.73-1.29).

DISCUSSION

The results of this IPD-MA showed that embryo selection with metabolomic profiling of culture medium by the NIR spectroscopy technique, as an adjunct to morphology does not improve live birth rates compared with embryo selection by morphology alone.
Our conclusion supports the general outcomes of three previously published RCTs[13-15]. In those RCTs, no beneficial effect on pregnancy and live birth rates of embryo selection by NIR plus morphology was found when compared with embryo selection by morphology alone. However, only one of those RCTs reported data of a completed study[14]. The publication by Hardarson et al.[13] showed data of an interim analysis and the publication of Sfontouris et al.[15] showed data of a prematurely terminated study due to the market withdrawal of the NIR instrument. Furthermore, in the interim analysis of the RCT by Hardarson et al.[13] minor, non-significant differences were reported in pregnancy outcomes between the two study groups for different days of transfer. Additionally, the RCT by Sfontouris et al.[15] did show significant differences in implantation rates in favour of the NIR group compared with the control group for day 5 transfers (but not for day 2 and day 3 transfers) yet no significant differences in ongoing pregnancy and live birth rates between the two study groups. By pooling the individual data of the three previously published RCTs and with the addition of data of another RCT, we increased the statistical power considerably. Even though the studies were heavily weighted by the two larger RCTs, the IPD-MA allowed us to standardize analyses across studies and explore patient- and study-level factors. Therefore, we were able to provide a more reliable estimate of treatment effect.
All of the authors who had conducted an RCT where the embryo selection was performed by morphology compared with embryo selection by morphology plus NIR technology provided their data. This avoided potential bias that could have occurred when investigators would not have been able to share original patient data. Important differences between the four IPD-MA studies were the day of transfer and the number of transferred embryos. Embryo transfer was performed on either day 2 (n=205), day 3 (n=512) or day 5 (n=207) and the two largest RCTs performed SET, while the two smaller studies generally transferred three embryos. The algorithms used varied depending on the day of transfer. In addition, the commercial instruments, used in the studies by Sfontouris et al.[15] and Economou et al., had the same set of pre-programmed algorithms, while the algorithms of the prototype instruments used in the studies by

Hardarson et al.[13] and Vergouw et al.[14] differed. Nevertheless, the data of the four studies showed no significant heterogeneity. Furthermore, in the multivariate logistic regression model, when adjusted for confounders such as day of transfer and number of transferred embryos, there was still no significant difference in live birth rates between the control group and the NIR group.

It has been proposed that a viable embryo changes its environment (culture media) differently from that of a non-viable embryo[18]. Therefore, we believe that the principle of assessing an embryo's reproductive potential by analysing culture medium constituents is still potentially valid. Recent examples that validate this hypothesis are the consistent use of glucose uptake as a marker of embryo viability in numerous species, including human[19] and the expression of soluble human leukocyte antigen G (sHLA-G) by embryos[20]. In this case, the used method (NIR spectroscopy technology in its current state) was not able to distinguish between viable and non-viable embryos more accurately than standard morphology. Instrument variability and susceptibility of the used algorithms to noise, which lowers the precision and repeatability of the NIR analyses, were possible reasons for the discrepancy between proof of principle studies and the larger RCTs[13]. In respect to this, the SET studies may have been pushing the technology's limits whereby the transfer of multiple embryos may have allowed more leeway for the lack of precision. The next step in spectroscopic assessment of culture media might be the evaluation of other spectroscopic techniques. Unfortunately, this strategy might turn out to be a rocky road, as Kirkegaard et al.[21] showed, in a prospective cohort study, large variations in individual spectra that did not correlate with pregnancy outcome using nuclear magnetic resonance (NMR) spectroscopy.

A frequently heard criticism of the NIR technology is its inability to identify specific culture metabolites. The NIR technology only allows overall spectral profile comparisons[22]. However, if the NIR technology had proven to be effective for embryo selection, this argument would have become less important. A common theme in all the RCTs was the differences in morphological status of the embryos. The NIR selected embryos all displayed poorer morphological criteria even though they were selected from a pool of 'good' embryos. The ability of these NIR selected embryos to provide the same pregnancy rates as the specifically selected 'best' morphology embryos provides some indication of an underlying benefit of the NIR instrumentation, or the inaccuracy of morphology to predict viability, i.e. to select the most viable embryo between good quality embryos.

The traditional embryo selection method by morphological criteria and cleavage rate has lately become a topic of debate among several experts. The reliance solely on a few predetermined static morphology evaluations and the considerable degree of intra-

and inter-observer variability in embryo grading among experienced embryologists[23,24] are the main concerns. The inability of morphological assessment to accurately predict the reproductive potential of an individual embryo has led to the search for new embryo selection tools. Several new technologies have been examined in search of additional markers of embryo viability to supplement the current criteria for embryo selection. Unfortunately, we already have experienced that novel techniques that have shown very promising results in retrospective analyses are often marketed for clinical use before proof of their effectiveness have been evaluated in RCTs. Even metabolomic profiling by NIR spectroscopy technology, the embryo selection method evaluated in the current IPD-MA, was marketed for routine IVF use after the proof of principle studies showed better pregnancy outcomes using this technology in adjunct to morphology. On a much greater scale, preimplantation genetic screening with the use of fluorescence in situ hybridization was already a frequently used technique in routine IVF laboratories before the promise of better pregnancy results was undermined by several RCTs[25-27]; summarized in ref. 28).

In the future, new technologies should be evaluated for effectiveness, safety and cost-effectiveness before they are introduced in routine IVF practice[29]. An important question to ask within this evaluation is: do couples for whom the new (embryo selection) technology is used have a better chance of having a baby than couples who do not[30]? Adequately powered and properly designed studies, preferably RCTs, are essential to determine whether it is the actual technology causing the effect or other factors. If appropriate, follow-up of children should be performed to analyse if the risks are within acceptable range[29]. The paradox, however, is that the companies trying to launch a product usually have limited funds to support large-scale and costly RCT studies.

In conclusion, there is at present no evidence that NIR spectroscopy of spent embryo culture media in its current form can be used in daily practice to improve live birth rates. To avoid the use of suboptimal embryo selection tools at the expense of the patient, evidence-based proof of clinical usefulness is essential before the implementation of new diagnostic tools in routine IVF practice.

REFERENCES

1. Kjelberg AT, Carlsson P, Bergh C. Randomized single versus double embryo transfer: obstetric and paediatric outcome and a cost-effectiveness analysis. Hum Reprod 2006; 21: 210-216.

2. Kallen B, Finnstrom O, Lindam A, Nilsson E, Nygren KG, Otterblad OP. Trends in delivery and neonatal outcome after in vitro fertilization in Sweden: data for 25 years. Hum Reprod 2010; 25: 1026–1034.

3. Bromer JG, Seli E. Assessment of embryo viability in assisted reproductive technology: shortcomings of current approaches and the emerging role of metabolomics. Curr Opin Obstet Gyneacol 2008; 20: 234-241.

4. Seli E, Sakkas D, Scott R, Kwok SC, Rosendhal SM, Burns DH. Noninvasive metabolomic profiling of embryo culture media using Raman and near-infrared spectroscopy correlates with reproductive potential of embryos in women undergoing in vitro fertilization. Fert Steril 2007; 88: 1350-1357.

5. Scott RT, Seli E, Miller K, Sakkas D, Scott K, Burns DH. Noninvasive metabolomic profiling of human embryo culture media using Raman spectroscopy predicts embryonic reproductive potential: a prospective blinded pilot study. Fert Steril 2008; 90: 77-83.

6. Vergouw CG, Botros LL, Roos P, Lens JW, Schats R, Hompes PGA, Burns DH, Lambalk CB. Metabolomic profiling by near infrared spectroscopy as a tool to assess embryo viability: a novel, non-invasive method for embryo selection. Hum Reprod 2008; 23: 1499-1504.

7. Nagy ZP, Sakkas D, Behr B. Non-invasive assessment of embryo viability by metabolomic profiling of culture media ('metabolomics'). Reprod Biomed Online 2008; 17: 502-507.

8. Nagy ZP, Jones-Colon S, Roos P, Botros L, Greco E, Dasig J, Behr B. Metabolomic assessment of oocyte viability. Reprod Biomed Online 2009; 18: 219-225.

9. Seli E, Vergouw CG, Morita H, Botros LL, Roos P, Lambalk CB, Yamashita N, Kato O, Sakkas D. Noninvasive metabolomic profiling as an adjunct to morphology for noninvasive embryo assessment in women undergoing single embryo transfer. Fert Steril 2010; 94: 535-542.

10. Ahlström A, Wikland M, Rogberg L, Barnett JS, Tucker M, Hardarson T. Cross-validation and predictive value of near-infrared spectroscopy algorithms for day-5 blastocyst transfer. Reprod Biomed Online 2011; 22: 477-484.

11. Seli E, Bruce C, Botros LL, Henson M, Roos P, Judge K, Hardarson T, Ahlström A, Harrison P, Henman M, Go K, Acevedo N, Siques J, Tucker M, Sakkas D. Receiver operating characteristic (ROC) analysis of day 5 morphology grading and metabolomic Viability Score on predicting implantation outcome. J Assist Reprod Genet 2011; 28: 137-144.

12. Vergouw CG, Botros LL, Judge K, Henson M, Roos P, Kostelijk EH, Schats R, Twisk JWR, Hompes PGA, Sakkas D, Lambalk CB. Non-invasive viability assessment of day 4 frozen-thawed human embryos by near-infrared spectroscopy. Reprod Biomed Online 2011; 23: 769-767.

13. Hardarson T, Ahlström A, Rogberg L, Botros L, Hillensjö T, Westlander G, Sakkas D, Wikland M. Non-invasive metabolomic profiling of day 2 and 5 embryo culture medium: a prospective randomised trial. Hum Reprod 2012; 27: 89-96.

14. Vergouw CG, Kieslinger DC, Kostelijk EH, Botros LL, Schats R, Hompes PGA, Sakkas D, Lambalk CB. Day 3 embryo selection by metabolomic profiling of embryo culture media with near-infrared spectroscopy as an adjunct to morphology: a randomized controlled trial. Hum Reprod 2012; 27: 2304-2311.

15. Sfontouris IA, Lainas GT, Sakkas D, Zorzovilis IZ, Petsas GK, Lainas TG. Non-invasive metabolomic analysis using a commercial NIR instrument for embryo selection. J Hum Reprod Sciences 2013; 6: 133-139.

16. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 1993; 341: 418-422.

17. Dutch Cochrane collaboration. Formulier II voor het beoordelen van een Randomised Controlled Trial (RCT). http://dcc.cochrane.org/sites/dcc.cochrane.org/files/uploads/RCT.pdf.

18. Leese HJ. Metabolism of the preimplantation embryo: 40 years on. Reproduction 2012; 143: 417-427.

19. Gardner DK, Wale PL, Collins R, Lane M. Glucose consumption of single post-compaction human embryos is predictive of embryo sex and live birth outcome. Hum Reprod 2012; 26: 1981-1986.

20. Kotze D, Kruger TF, Lombard C, Padayachee T, Keskintepe L, Sher G. The effect of the biochemical marker soluble human leukocyte antigen G on pregnancy outcome in assisted reproductive technology-a multi center study. Fert Steril 2013; 100: 1303-1309.

21. Kirkegaard K, Svane ASP, Hindkjær JJ, Nielsen NC, Ingerslev HJ. Metabolic profiles of spent culture media does not predict pregnancy outcome. Hum Reprod 2013; 28(Suppl 1): i23-i25.

22. Botros LL, Sakkas D, Seli E. Metabolomics and its application for non-invasive embryo assessment in IVF. Mol Hum Reprod 2008; 14: 679-690.

23. Baxter Bendus AE, Mayer JF, Shipley SK, Catherino WH. Interobserver and intraobserver variation in day 3 embryo grading. Fert Steril 2006; 86: 1608-1615.

24. Paternot G, Devroe J, Debrock S, D'Hooghe T, Spiessens C. Intra- and inter-observer analysis in the morphological assessment of early-stage embryos. Reprod Biol Endocrin 2009; 7: 105-10.

25. Staessen C, Verpoest W, Donoso P, Haentjens P, van der Elst J, Liebaers I, Devroey P. Preimplantation genetic screening does not improve delivery rate in women under the age of 36 following single-embryo transfer. Hum Reprod 2008; 23: 2818-2825.

26. Mastenbroek S, Twisk M, van Echten-Arends J, Sikkema-Raddatz B, Korevaar JC, Verhoeve HR, Vogel NEA, Arts EGMJ, de Vries JWA, Bossuyt PM, Buys CHCM, Heineman MJ, Repping S, van der Veen F. In Vitro Fertilization with Preimplantation Genetic Screening. N Engl J Med 2007; 357: 9-17.

27. Hardarson T, Hanson C, Lundin K, Hillensjö T, Nilsson L, Stevic J, Reismer E, Borg K, Wikland M, Bergh C. Preimplantation genetic screening in women of advanced maternal age caused a decrease in clinical pregnancy rate: a randomized controlled trial. Hum Reprod 2008; 23: 2806-2812.

28. Mastenbroek S, Twisk M, van der Veen F, Repping S. Preimplantation genetic screening: a systematic review and meta-analysis of RCTs. Hum Reprod Update 2011; 17: 454-466.

29. Harper J, Magli MC, Lundin K, Barratt CLR, Brison D. When and how should new technology be introduced into the IVF laboratory? Hum Reprod 2012; 27: 303-313.

30. Van Steirteghem A. What next for assisted reproductive technology? A plea for an evidence based approach. Hum Reprod 2010; 23: 2615-2616.